An experimental analysis of experiential and cognitive variables in web navigation

Paul van Schaik

Teesside University

Jonathan Ling

University of Sunderland

To appear in Human-Computer Interaction

RUNNING HEAD: A MODEL OF WEB NAVIGATION

Corresponding Author's Contact Information:

School of Social Sciences and Law
Teesside University
Borough Road
Middlesbrough, Cleveland, TS1 3BA
United Kingdom
Tel. +44 1642 342320
Fax +44 1642 342399
p.van-schaik@tees.ac.uk

Brief Authors' Biographies:

Paul van Schaik is a psychologist with an interest in human-computer interaction, and cognitive psychology; he is a Professor of Psychology in the School of Social Sciences and Law at Teesside University. Jonathan Ling is a psychologist with an interest in human-computer interaction, cognitive performance and perception; he is a research fellow in the Faculty of Applied Sciences at the University of Sunderland.

Abstract

Cognitive and experiential factors in human-computer interaction have been the focus of significant recent attention, but there is a lack of a much needed integrated approach to these issues. The current paper proposes such an approach and applies this, combined with the person-task-artefact model (Finneran & Zhang, 2003), to the modelling of web navigation. In an experiment, artefact complexity and task complexity were manipulated. The effects of the experimental manipulations and intrinsic motivation on flow experience, task performance and task outcome were tested. The main effects of the manipulations were confirmed. Further analyses demonstrated that flow was a mediator of the effect of experimental manipulations on task performance, and task performance was a mediator of the effect of flow on task outcome. Overall, the results in the domain of web navigation that are presented here demonstrate the need for taking an integrated cognitive-experiential approach in the modelling of human-computer interaction.

Contents

1. Background

"I can never think and play at the same time. It's emotionally impossible."

– From *The New Tristano* (Lennie Tristano, 1962)

Although people's conscious reflection on their experience can interfere with highly involved and accomplished task performance (such as improvising and simultaneously accompanying oneself in a structured piece of music such as Tristano's *C Minor Complex*), the main argument of this paper is that a positive interaction experience enhances task performance and thereby, indirectly, the outcome of such performance. Cognitive factors have been studied in human-computer interaction and related disciplines, such as ergonomics, for many years (see, e.g., Broadbent, 1958), particularly in relation to task performance and usability. For instance, the study of usability has been approached using a cognitive psychological framework (e.g. Gardiner & Christie, 1987). Models of cognitive task performance have been developed to analyse and aid the design of interactive computer systems. Examples include Fitt's law, Hick's law and the Goals Operators Methods and Selection rules (GOMS) model, based on the Model Human Information Processor (Card, Moran & Newell, 1983). The study of experiential factors in human-computer interaction has started more recently (e.g. Hassenzahl & Tractinsky, 2006; Law & van Schaik, 2010). Research on interaction experience ('user experience') has used various social-psychology and other approaches (e.g. Hassenzahl & Monk, in press). Moreover, models of interaction experience have been proposed (e.g. Hassenzahl, 2004) and tested (e.g. van Schaik & Ling, 2008).

Thus, cognitive and experiential factors of human-computer interaction have been studied separately, but often not in an integrated fashion. However, there is evidence that cognitive-task performance is influenced by both cognitive and experiential factors. For example, Moshagen, Musch and Göritz (2009) demonstrated that a visual design with enhanced aesthetics can improve cognitive-task performance using a Web site under conditions of poor usability. These results are further strengthened by similar findings reported by Sonderegger and Sauer (2010). Moshagen et al. invoke Norman's (2004) positive-affect-mediation model as an account for their findings in which "positive affect helps to overcome the obstacles posed by difficult problems" (Moshagen et al., p. 1316). However, according to Hasssenzahl (2008), Norman's model cannot explain the effect of positive affect in all situations. Based on existing social-psychology theory and empirical evidence, Hassenzahl makes different predictions of the effect of positive affect, depending on the type of task that a user performs. Furthermore, in domains other than human-computer interaction, flow experience has been found to predict cognitive performance over and above existing skills and knowledge (Engeser & Rheinberg, 2008). Moreover, evidence from the study of neurological patients demonstrates that cognitive performance is impossible without emotional experience (Damasio, 2000). In addition, within the framework of the 'information seek cycle', David et al.'s (2007) experiential dynamic modelling of web navigation should also be considered. They showed that self-efficacy is enhanced by the successful execution of information-seeking goals in one cycle. This, in turn, reduces the perceived difficulty of information goals in the following cycle. In addition, as a result of self-efficacy from previous cycles, more challenging goals are formulated in subsequent cycles. Although the effect on cognitive task performance was not studied, given the nature of David et al.'s 'virtuous circle', it would be expected that such a circle would produce enhanced performance. Given the evidence from these studies, there is a need for an integrated approach to studying cognitive and experiential factors in human-computer interaction. Here we apply this approach to the modelling of web navigation, with a focus on flow experience – a ‘holistic sensation that people feel when they act with total involvement’ (Csikszentmihalyi, 1990, p. 477).

Different approaches to the study of cognitive and experiential factors in human-computer interaction have been reported. In order to understand the differences it is useful to contrast two approaches: one stressing the uniqueness of individual experiences and another stressing the commonalities between people in their experiences – while still allowing for variation depending on, for example, task characteristics and individual differences. A simple, but extremely useful, example given by Hassenzahl (2010, p. 74) clarifies the distinction in approach.

> Take marriage counselling as an example. I’m sure that the problems couples encounter and the actual experiences they make in their relationships are rich and diverse. A thorough understanding of each case is necessary to pick or even develop appropriate approaches to, for example, solve marital problems. However, counselling and intervention also requires a more general understanding of what intimate relationships have in common and how classes of marital problems should or shouldn’t be approached. Both is a reduction—the combined result of many people trying to understand this problem domain, the distillate of hours and hours of counselling work. No seasoned practitioner or researcher would dismiss this knowledge, and always start afresh each time.

The distinction between an emphasis on uniqueness and commonality goes back to a more fundamental and profound one that is found in psychology, the science of human behaviour and mental activity, regarding the nature of human behaviour (in general, irrespective of whether technology is used or not). Given that a comprehensive review of all research paradigms in psychology is beyond the scope of this publication, two illustrative examples are presented here to clarify the distinction – phenomenology (stressing uniqueness) and the information- processing approach (stressing commonality) – and applied to human-computer interaction.

The origins of phenomenology date back to Edmund Husserl (1859-1938). Although different perspectives on phenomenology have been advanced, they have in common that they are concerned with the study of the 'lived experiences' of people, they consider these experiences to be of a conscious nature and they are concerned with the development of the description of the 'essences' of these experiences rather than with explanations or analyses (Creswell, 2007). Applications of phenomenology are for example found in cognitive psychology in the study of autobiographical memory (Conway, 2002), in health psychology in the study of patients' experience of their physical illness (Ayers & Forshaw, 2010) and in counselling psychology in the study of therapists' experience of their counselling practice (Rizq & Target, 2008). Phenomenology has been applied to the study of human-computer interaction to capture the essence of individual instances of interaction experience (e.g. McCarthy and Wright, 2004). This work has not aimed at developing a (family of) more generally applicable models or categories of users' interaction experience as proposed by other research (e.g. Hassenzahl et al., 2010; van Schaik & Ling, in press; Sutcliffe, 2010), but ostensibly dismissed the attractive prospect of using existing psychological theory that is distilled in numerous relevant theoretical and empirical research publications. In addition, Hassenzahl et al. (2010, p. 354) have noted another problem, related to design, with this approach: "... in the moment of description, [experiences] are gone and will never occur again. This actually would be the end of story for experience in HCI, because designing for bygone and unrepeatable experiences is futile".

The information-processing approach in psychology started with the Symposium on Information Theory, held at the Massachusetts Institute of Technology in 1956. In this approach, people are studied as information processors, in particular in terms of representations and processes that operate on the representations. Research aims to identify these processes and representations by inference from human behaviour. Since 1956 this approach has been fruitfully applied in psychology and allied disciplines such as management science (e.g. Gefen, Karahanna & Straub, 2003) and behavioural economics (e.g. Thaler & Sunstein, 2009). Within psychology, the approach has been applied to study human behaviour at the level of biological processes (e.g. gait-control theory of pain; Melzack & Wall, 1965), cognitive processes (e.g. Kintsch's [1998] construction-integration theory of reading) and social cognition (e.g. Bandura's [1986] social cognitive theory). In human-computer interaction the information-processing approach has also been fruitfully applied. Examples include Card, Moran and Newell's (1983) model human information processor, Kitajima and Blackmon's cognitive model of web navigation (e.g. Blackmon, Kitajima & Polson, 2005) and David et al.'s (2007) information-seek-cycle.

In this paper, we take an information-processing approach. This is because both more than half a century of refereed published research and people's common experience demonstrate that it is possible to explain and predict human behaviour, in general and in particular in the case of using artefacts, to a significant extent (but not perfectly), as a consequence of commonalities between people in terms of not only their cognitive processes, but also their experience (Sheldon, Elliot, Kim & Kasser, 2001). A perfect account is not possible because of variation due to task characteristics and individual differences in ability and personality. In relation to human-computer interaction, a useful consequence of the information-processing approach is that it offers the prospect of design guidance that can improve system design for human users in terms of psychologically grounded generic classes of experience (Hassenzahl, Diefenbach & Göritz, 2010), but not necessarily a perfect design because of unexplained variation in human behaviour (for example due to individual differences between users). The information-processing approach is also useful as it facilitates the experimental study of human-computer interaction in the framework of the Finneran and Zhang's (2003) person-artefact-task model (introduced in the next section). This can be done by experimentally manipulating the effects of artefact and task variables, and measuring individual-difference variables (e.g. spatial ability) and then observing their effect on cognitive measures (for example, in terms of work load) and experiential variables (for example, flow experience, introduced in the next section) as well as testing predictions about the relations between cognitive and experiential variables. In this way, research contributes to the knowledge base of human-computer interaction, and the current study adopts this modus operandi.

2. Introduction

Although interactive-system characteristics (e.g. compliance with cognitive design principles; Dalal et al. 2000) are an important factor in users' success in finding information in web sites; other types of factor are also influential. We propose using Finneran and Zhang's (2003) person-artefact-task model (see Figure 1) – originally developed for modelling flow experience in human-computer interaction – as a framework for modelling people's 'navigation' of web sites, where a computer user follows a path through a web site by visiting web pages in order to complete a task. Using this framework, characteristics of an *artefact* (e.g. the complexity of a web site), the *task* performed with the artefact (e.g. the complexity of a task) and those of the *person* (end-user) performing the task (e.g. a user's spatial ability) all influence the process of task performance (e.g. the extent to which users can orientate themselves in a web site). Finneran and Zhang note that the effects of person, artefact and task on the process are not independent. Therefore, it is important to consider not only the effects of these three entities separately, but also the ways in which the effect of each may depend on the influence of the others. Finally, person, task and artefact – through the process of task performance – influence task outcomes (e.g. the success rate of finding information in a particular web site, given a particular information-retrieval task).

__________ Insert Figure 1 about here. __________

The process of task performance comprises components of experience and cognition, and can be further analysed using the concept of flow experience ('flow' for short). Consistent with Finneran and Zhang's (2003) original person-artefact-task model, we studied flow separately from other aspects of task process (e.g. disorientation). Flow is defined as a 'holistic sensation that people feel when they act with total involvement' (Csikszentmihalyi, 1990, p. 477). Nine dimensions of flow (see Figure 2) have been distinguished and measurement instruments for these dimensions have been developed and validated (e.g. Jackson & Marsh, 1996). It is important to note that flow is not a matter of 'all or nothing', but people can experience a degree of flow on each of the dimensions. In the context of human-computer interaction, using a bottom-up research approach (the method of grounded theory), Pace (2004) provides further evidence for the dimensions of flow in web users. Furthermore, Guo and Poole (2009) measured the nine dimensions of flow experience in human-computer interaction. By conceptualising and measuring flow more comprehensively, these two studies represent an advance over previous work that has studied flow experience in human-computer interaction using fewer dimensions (Davis & Wiedenbeck, 2000; Novak, Hoffman & Duhachek, 2003).

In the context of web navigation, the effects of *artefact* characteristics, such as link colour (e.g. Pearson & van Schaik, 2003) and screen layout (e.g. van Schaik & Ling, 2006), on various measures has been investigated. Here, we study the complexity of web site design as an artefact characteristic, as recent research has identified this complexity as a factor that increases task difficulty. According to Mosenthal (1996), documents in general and, by implication, web sites, in particular differ, first, in their complexity, that is the number of information items and labels in, and the structural organisation of a document and, second, in the 'plausability of distractors', that is the extent to which competing information items partially, but not fully, match the information request. Mosenthal found that both of these are significant predictors of reading-task difficulty. According to more recent research, page complexity in terms of the number of navigation choices on a web page (Gwidzka & Spence, 2006) and structural complexity (Guo & Poole, 2009) may increase task difficulty. Guo and Poole found indeed that perceived artefact complexity decreases flow experience. As further empirical evidence for the negative effect of complexity, Blackmon, Kitajima, Polson and Lewis (2002) reported that the greater the number of links per page the lower the success rate (in terms of percentage correct [links clicked] on first click).

Theoretical support for the negative effect of artefact complexity may seem to come from Hick's law: task difficulty increases with the number of response alternatives (Usher et al., 2002), but this only applies to tasks involving a restricted type of stimulus, highly compatible stimuli and responses, and responses that are well practised (Paap & Cooke, 1997). However, the dual-criterion model of menu selection (Pierce, Sisson & Parkinson, 1992) – which is based on semantic-memory models – is consistent with a decrease in accuracy of option (link) selection with an increasing number of options (links). This model distinguishes four distributions of the likelihood that a target belongs to the category described by a (menu) option for (a) targets that are typical members of the category, (b) targets that are atypical members, (c) targets that are non-members of the category, but related to the category and (d) targets that are non-members and not related. Two criteria are distinguished: a high criterion (H) and a low criterion (L). Comparisons producing a value below L will immediately be rejected and those producing a value above H will immediately be selected, leading to a self-terminating search. Comparisons with values between L and H will produce candidates that are considered in case the search was not self-terminating. If the comparison process produces one candidate then the corresponding option will be chosen. If there are more candidates then (a subset of) these will be re-examined and a partially redundant search occurs. The mechanism for producing a larger set of candidates with an increasing number of options is that the value of H decreases. This larger set of competing candidates increases the chance of incorrect selection. Furthermore, as artefact complexity increases, the balance between challenge and skill will be negatively affected and flow experience will decrease (Guo & Poole, 2009). Therefore,

Hypothesis 1a: artefact complexity (page complexity) has a negative effect on task performance.

Hypothesis 1b: artefact complexity (page complexity) has a negative effect on flow experience.

Hypothesis 1c: artefact complexity (page complexity) has a negative effect on task outcome.

Tasks involved in web navigation also differ in several respects. For example, information-retrieval tasks and learning tasks have been distinguished (Rouet, 2006). Within the former, tasks still vary for example in the information that is being retrieved. Tasks may differ in terms of the type of information (i.e. concrete or abstract) that is being requested, and the type of match between the information requested and the actual available information, both of which have been shown to be significant predictors of reading-task difficulty (Mosenthal, 1996). Here, we study an aspect of web navigation tasks – task complexity – which has been identified as a factor that can decrease the quality of human-computer interaction (Gwizdka & Spence, 2006). Van Oostendorp, Madrid and Puerta Melguizo's (2009) findings show that task complexity (path length, defined as the number of steps involved in finding the information) has a negative effect on task performance. Based on Gwizdka and Spence (2006), two mechanisms for this effect can be distinguished. First, for a particular task goal and a given probability of selecting the correct link on each page on the path to the page containing the target information, it follows that the longer the path to the target, the smaller the probability of following a path consisting of only correct links. In other words, the longer the task path, the greater the chance of making at least one error and following an incorrect path. Second, a similar argument goes for the probability of making correct relevance judgements about the information presented in each of a sequence of web pages along the path: the longer the path, the greater the chance of making an incorrect relevance judgement either by judging a relevant piece of information to be irrelevant for the task goal that is being pursued or vice versa. Therefore, task performance will decrease with path length. In addition, as task complexity increases the balance between challenge and skill will be adversely affected and flow experience will decrease (Guo & Poole, 2009). Thus,

Hypothesis 2a: task complexity (path length) has a negative effect on task performance.

Hypothesis 2b: task complexity (path length) has a negative effect on flow experience.

Hypothesis 2c: task complexity (path length) has a negative effect on task outcome.

Individual-difference variables in *persons* (end-users) have an effect on their navigation of web-based systems. For example, Juvina and van Oostendorp (2006) demonstrated that spatial ability and domain expertise are positive predictors of task performance (in terms of effectiveness defined as a combination of correctness and completeness), but working-memory capacity is a negative predictor of disorientation. Here we study intrinsic motivation as a person-characteristic (individual difference variable) in web navigation. Intrinsic motivation, in this sense, is a positive predictor of task outcome in academic learning (Hirschfeld et al., 2008) and school learning (Vansteenkiste et al., 2008), and of flow experience in golf (Oh, 2001) and athletics (Stavrou, 2008). Theoretically, there are several reasons for a positive effect of intrinsic motivation on task performance (Zapata-Phelan et al., 2009). First, activity, concentration, initiative, resilience and flexibility can enhance task performance. Second, intrinsic motivation has a stronger effect than external motivation on the persistence of effort, which has a strong effect on the performance of complex tasks or complex artefacts, but this effect may be weaker or disappear with less complex tasks and less complex artefacts. Third, in the domain of employment, internal (work) motivation is expected to have a positive effect on task performance. Intrinsically motivated individuals (individuals with an ‘autotelic’ personality) are those who engage in activities for the sake of the activities rather than in order to achieve some external goal. Therefore, these individuals should experience a higher level of well-being (Asakawa, 2004, 2009) and a higher level of flow than others (Asakawa, 2004). Therefore, *Hypothesis 3a*: intrinsic motivation has a positive effect on task performance.

Hypothesis 3b: intrinsic motivation has a positive effect on flow experience.

Hypothesis 3c: intrinsic motivation has a positive effect on task outcome.

It has been demonstrated that flow experience is an *independent* positive predictor of task outcome (after controlling for other [cognitive] variables) in the domains of computer-game playing (Murphy, Smith & Hancock, 2008), mathematics (Heine, 1997; Engeser & Rheinberg, 2008), a foreign language (Engeser & Rheinberg, 2008) and computer-based statistics (Vollmeyer & Imhof, 2007). Engeser and Rheinberg (2008) propose two pathways for the positive effect of flow on performance outcome. First, flow is considered to be a 'highly functional state' (where 'individuals ... are highly concentrated and optimally challenged while being in control of the action' [Engeser & Rheinberg, 2008, p. 158]) which, therefore, should promote performance. Second, flow is a driver of motivation for continued activity, which leads people to select higher challenges in order to experience flow again. Thus,

Hypothesis 4: flow experience has a positive effect on task outcome, after controlling for the influence of artefact complexity, task complexity and intrinsic motivation.

__________ Insert Figure 2 about here. __________

Flow initially develops as a result of performing a task. The experience of flow is an intrinsic-motivation construct (Schwartz & Waterman, 2006) and can develop during an activity as a "deep involvement in an activity that is perceived as rewarding in and of itself" (Keller & Bless, 2008, p. 196). Consequently, flow can act as a motivating force to continue task performance. It is important to note that this type of (state) motivation develops during task performance as a result of the experience of task performance. In contrast, intrinsic (trait) motivation as an individual-difference variable is considered as a stable (personality) characteristic. Given the potential role of flow as a motivator, flow has a positive effect on task performance, which – in turn – has a positive effect on task outcome. Furthermore, although it has been both theoretically and empirically confirmed that flow is an antecedent of task performance outcome, flow – as an experience – cannot in and of itself directly influence the outcome. Instead, task outcome is achieved by and is the result of performing a task. Thus, the higher the quality of this performance, the higher the quality of task outcome. Hence, the positive effect of flow on task outcome must be mediated. Given the motivating character of flow to continue task performance, the quality of task performance is a likely mediator. Accordingly, flow experience should have a positive effect on task performance and, through this route, a positive (indirect) effect on task outcome. Therefore,

Hypothesis 5: flow experience has a positive effect on task performance, after controlling for artefact complexity, task complexity and intrinsic motivation.

Hypothesis 6: task performance has a positive effect on task outcome, after controlling for artefact complexity, task complexity, intrinsic motivation and flow.

The aim of the current study is to demonstrate the need for an integrated approach to studying cognitive and experiential factors in human-computer interaction in order to model web navigation. We do this by testing the hypotheses using a computer-controlled experiment, in which artefact complexity and task complexity were manipulated. Test-users' intrinsic motivation was measured as an individual-difference variable. In a series of information retrieval tasks, users employed an information-oriented realistic mock intranet site.

3. Method

3.1. Design

A 2×2 between-subjects experimental design was used with the independent variables of artefact complexity (high and low) and task complexity (high and low). (Relatively) low artefact complexity was defined as five links and (relatively) high complexity as ten links on each web page. In a low-complexity task the answer was available on a page one or two links from the homepage. In a high-complexity task the answer was available on a page four links from the homepage. In addition, the uncontrolled individual-difference variable of intrinsic motivation was measured (by questionnaire; see Section 3.3) to establish its effect on flow.

Dependent variables included both in-task measures (of mental effort, task performance and task outcome) – as well as retrospective measures (of flow and disorientation). Task performance was measured in terms of efficiency – in terms of time-on-task on correctly completed tasks, number of completed tasks and disorientation (as an indicator measure of the efficiency of navigation behaviour; see van Oostendorp et al., 2009) and task difficulty – in terms of perceived workload. Flow experience was measured by questionnaire (see Section 2.3). Task outcome was measured as the correctness of reported search result/answer in terms of percentage of correct answers relative to the total number of tasks and percentage of correctly completed tasks.

3.2. Participants

One hundred and fourteen undergraduate psychology students (91 females and 23 males), with a mean age of 22.66 years (*SD* = 6.03) took part in the experiment as a course requirement. There were 30 participants in the condition of low artefact complexity/low task complexity, 29 in the low/high condition, 28 in the high/low condition and 27 in the high/high condition. All participants had used the Web. Mean experience using the Web was 9.68 years (*SD* = 3.03), mean time per week spent using the Web was 17.25 hr (*SD* = 16.73) and mean frequency of Web use per week was 14.76 (*SD* = 9.87). This type of sample was necessary for the specific site (a psychology intranet) that was studied and would be typical of users of these intranets – in terms of studying psychology and being predominantly female, but not necessarily of intranet sites or Web sites in general. As participants took part as a course requirement without additional reward, neither specific intrinsically motivating factors nor specific extrinsically motivating factors that would influence the results were likely.

3.3. Materials and equipment

Participants gave responses to several questionnaires, using 7-point scales (see Appendix), except where indicated otherwise. Four items from Guay, Vallerand and Blanchard's (2000) 16-item Situational Motivation Scale (SIMS) measured intrinsic motivation. In order to measure the dependent variables, several measurement scales were used. Zijlstra's (1993) single-item Subjective Mental Effort Questionnaire (SMEQ) ranging from 0 ('hardly effortful') to 220 ('exceptionally effortful') was employed to measure mental effort. The nine dimensions of flow experience were measured using Jackson and Marsh's (1996) 36-item Flow State Scale (FSS), as this instrument measures flow more comprehensively than other instruments (e.g. Davis & Wiedenbeck, 2000; Novak, Hoffman & Duhachek, 2003). The psychometric properties of the nine dimensions are reported in Section 4.1, but to make the testing of hypotheses tractable a composite measure is used in Sections 4.2 and 4.3. Disorientation was measured with Ahuja and Webster's (2001) 7-item disorientation scale.

Two versions of a Web site were modelled as a typical psychology site for university students, and especially designed and programmed for the experiment. In addition to the homepage, the main pages of the high-complexity version (see Figure 3) of the site were Teaching, Research, Fees and Funding, Hall of Fame, Library, Staff, Sports and Leisure, Careers and About, with 9990 further Web pages and links to Web pages. In addition to the home page, the main pages of the low-complexity version (see Figure 3) of the site were Teaching, Research, Fees and Funding and Hall of Fame, with 620 further Web pages and links to Web pages. All links and content of the low-complexity version were also included in the high-complexity version. The complex web site had more pages than the simple web site, but both sites had an equal number of levels of depth (from the home page) and therefore both site versions allowed simple and complex tasks to be completed. The experiment ran on personal computers (Intel Pentium, 1.86 GHz, 2 GB RAM, Microsoft Windows XP operating system, 17-inch monitors). The screen dimensions were 1280×1024. Contrast (50%) and brightness (75%) were set to optimal levels.

__________ Insert Figure 3 about here. __________

3.4. Procedure

In Phase 1, participants completed the SIMS. In Phase 2, an information retrieval task followed which included typical tasks that users perform with educational intranet sites. In each trial, a question appeared at the top of the screen, for instance ‘How long is the exam for the module Introduction to Personality and Social Psychology?’ Once participants had read the question, they had to click on a button labelled 'Show web site'. The home page of the site then appeared on the screen and they had to find the answer to the question using the site. Participants were told to take the most direct route possible to locate the answer. Having found this, they clicked on a button labelled 'Your answer', which opened a dialogue box at the bottom of the screen. Participants typed their answers into the box, clicked on 'OK', completed the SMEQ for the task they had just performed and moved on to the next question. After three practice questions, the main set of information retrieval tasks followed, with a duration of 20 minutes – in which a maximum of 37 further questions were presented. After the information retrieval task, in Phase 3 participants completed the FSS and the disorientation scale. Finally, participants answered questions requesting demographic details. The experiment took about 35 minutes to complete.

3.5. Data analysis

Partial-least-squares path modelling (PLS for short; Vinzi, Chin, Henseler, & Wang, 2010) was used for data analysis for the following reasons (see also Streukens, Wetzels, Daryanto & de Ruyter [2010] for an example of the application and further justification for the use of PLS in experimental research). PLS allows the analysis of (single-stage and) multi-stage models with latent variables, allowing the integrated analysis of a measurement model and a structural model. Each latent variable (usually a psychological construct) is measured using one or more manifest variables (usually psychometric items). In contrast to covariance-based structural equation modelling techniques, PLS explicitly supports both reflective measurement and formative measurement, which is needed in research using both psychometric measurement and experimental manipulation of independent variables. PLS does not require some of the assumptions imposed by covariance-based structural equation modelling – including those of large sample size, and univariate and multivariate normality. Recent simulation studies have demonstrated that PLS path modelling performs at least as well as and under various conditions is superior to covariance-based structural equation modelling in terms of bias, root mean square error and mean absolute deviation (Hulland et al., 2010; Vilares et al., 2010). PLS is compatible with multiple regression analysis, analysis of variance and unrelated t tests, the results of which are special cases of the results of PLS. For an integrated and consistent approach, all analyses were conducted using PLS by way of the SmartPLS software (Ringle, Wende & Will, 2005), unless stated otherwise. In PLS analyses, a bootstrapping procedure (N = 5000, as Henseler, Ringle & Sinkovics, 2009, recommend) was used to test the significance of model parameters.

4. Results

First, the findings regarding psychometric properties of the measurement instruments are presented to establish the quality of measurement and in particular demonstrate that the structure of flow, identified in other domains, also applies to web navigation. Then, descriptive statistics and effect size are presented to show the effects of the experimental manipulations on the dependent variables and inferential statistics are presented to test the proposed hypotheses.

4.1. Psychometric properties of measurement instruments

The results presented here are those from the analysis of the effect of the independent variables and intrinsic motivation on the dependent variables. This is because this analysis includes all the psychometrically measured latent variables (flow, disorientation and intrinsic motivation). In testing the measurement model, reliability was analysed (see Figure 4), and convergent and discriminant validity was assessed (see Figure 5). The reliability of each individual reflective item is assessed by its loading on the construct of which it is an indicator, which should be 0.7 or higher (Henseler et al., 2009). After removing three items from the FSS (Items 1 [balance of challenge and skill], 33 [control] and 34 [loss of self-consciousness]), all the loadings exceeded this cut-off point. Using a bootstrapping procedure, the loadings of all items were found to be statistically significant. At the construct level, reliability was analysed using the composite reliability co-efficient, which needs to be 0.7 or higher. All the co-efficients exceeded this cut-off point.

__________ Insert Figure 4 about here. __________

__________ Insert Figure 5 about here. __________

Convergent validity (the extent of consistency among the items measuring a particular construct) was analysed using the average variance extracted (AVE) by a construct from its indicators, which should be 0.7 or higher (Henseler et al., 2009). All values exceeded this cut-off point. Discriminant validity (the extent to which a measure of a particular construct differs from measures of other constructs) was assessed by analysing the square root of the AVE by each construct from its indicators, which – according to the Fornell-Larcker-criterion – should be greater than its correlation with the remaining constructs. All values met this condition. In conclusion, the reliability, and the convergent and discriminant validity of the multi-item constructs was confirmed. The findings of discriminant validity provide evidence for the distinction between the dimensions of flow as a differentiated set of indicators of experience and both intrinsic motivation as a trait and disorientation as an indicator of task performance. Based on Cohen's (1988) conventions for effect size of r (0.1 for small, 0.3 for medium and 0.5 for large), the medium-sized correlation between intrinsic motivation and autotelic experience provides evidence for the conjecture that intrinsic motivation has a positive effect on this particular dimension of flow. The evidence gains further strength, as this correlation remained statistically significant, $t = 6.47$, $p < 0.001$, after controlling for the independent variables artefact complexity and task complexity, their interaction and the interaction of intrinsic motivation with each of these variables and their interaction. In order to test the hypotheses regarding flow, in the following analyses the measurement of flow was reduced to a composite measure created from five factors (mergence of action and awareness, clarity of goals, balance of challenge and skills, control and feedback) as they loaded highly on a single higher-order flow factor. The composite score was created from the scores on each of the five factors, using the PLS weighted-average algorithm.

4.2. The effects of artefact complexity, task complexity and intrinsic motivation on outcome measures

Descriptives showing outcome measures as a function of task complexity and artefact complexity are presented in Figure 6. In order to test *Hypotheses 1a-3c*, PLS analysis was conducted regarding the effect of task- and artefact complexity, and intrinsic motivation on task performance, flow and task outcome as dependent variables (see Figure 7). The negative effect of task complexity on objective measures of task performance and task outcome (task completion, time-on-task and correctness) was large or medium-sized (see Figure 6). The negative effect of artefact complexity was medium-sized. The interaction effect on correctly completed questions, where the negative effect of task complexity was stronger with high artefact complexity and vice versa, was small to medium-sized. The effects of task complexity and artefact complexity were significant (except for the effect of task complexity on time-on-task), in support of *Hypotheses 1a* and *1c*, and Hypotheses *2a* and *2c* (see Figure 7). The two effects were independent, except for correctness of completed tasks, where the negative effect of task complexity was stronger with high artefact complexity and vice versa. The negative effect of task complexity on subjective measures of task performance (work load and disorientation) and flow experience was medium-sized (see Figure 6). The negative effect of artefact complexity was (small to) medium-sized. The effects of task complexity and artefact complexity on task performance were significant, support in Hypotheses 1a and 2a.

The effect of task complexity on flow was also significant, supporting Hypothesis 2b; however, the effect of artefact complexity was not, so Hypothesis 1b was not supported. The results show no evidence for *Hypotheses 3a, 3b and 3c*, as the effect of intrinsic motivation on task performance, flow and task outcome was not significant.

__________ Insert Figure 6 about here __________

__________ Insert Figure 7 about here __________

4.3. The effects of task performance and flow experience on task outcome

In order to test *Hypotheses 4-6*, PLS analyses were conducted with artefact complexity, task complexity, task performance and flow experience as independent variables and task outcome (percentage of correctly completed tasks) as dependent variables (see Figure 8). The task performance measures were workload and disorientation; number of tasks completed was not used, as the outcome measure (correctly completed tasks) controls for this; time-on-task was not used, as this measure was confounded with task complexity. Intrinsic motivation was not used as a predictor here because in the test results of *Hypotheses 3a-3c* it was not a predictor of task outcomes.

__________ Insert Figure 8 about here __________

In PLS analyses, the contribution of the experimental manipulations (artefact complexity and task complexity combined) and two components of task process – task performance (disorientation and workload combined) and flow experience (a composite of action-awareness merging, autotelic experience, clarity of goals, balance of challenge and skills, control and feedback) – in explaining variance in correctness was analysed. Technically, predicted scores on the dependent variable from three sets of predictors (experimental manipulations, flow experience and task performance), simultaneously predicting the dependent variable, were used. This approach was followed to produce straightforward tests of *Hypotheses 4-6*. In accordance with these hypotheses, of interest were the effects of the combined experimental manipulations (including their interaction) and task performance rather than separate elements of these. The combined predictors explained 65% of variance in *correctness of completed tasks*.

In order to test *Hypothesis 4*, correctness was regressed onto flow and experimental manipulations (see Figure 8a). In support of *Hypothesis 4*, it was found that flow remained a significant predictor of correctness after controlling for experimental manipulations. In order to test *Hypothesis 5*, task performance was regressed onto flow and experimental manipulations (see Figure 8b). In support of *Hypothesis 5*, it was found that flow remained a significant predictor of task performance after controlling for experimental manipulations. In addition, experimental manipulations were a significant predictor of flow; flow was a significant predictor of task performance without controlling for experimental manipulations; and experimental manipulations was a significant predictor of task performance without controlling for flow. Given the test results for *Hypothesis 5* and the additional results, flow was a partial mediator (as the effect of experimental manipulations remained significant after controlling for flow) of experimental manipulations on task performance. In order to test *Hypothesis 6*, correctness was regressed onto task performance, flow and experimental manipulations. In support of *Hypothesis 6*, it was found that task performance remained a significant predictor of correctness after controlling for experimental manipulations and flow. In addition, flow and experimental manipulations together were significant predictors of task performance (see Figure 8b); task performance was a significant predictor of correctness without controlling for experimental manipulations and flow; and flow and experimental manipulations together were significant predictors of correctness (see Figure 8a). Given the test results for *Hypothesis 6* and the additional results, task performance was a full mediator (as the effect of flow on correctness became non-significant after controlling for task performance and experimental manipulations) of the effect of task performance and correctness.

5. Discussion

5.1. Exploration of main findings

The aim of the current study was to demonstrate the need for an integrated approach to studying cognitive and experiential factors in human-computer interaction to modelling web navigation by testing a set of hypotheses in an experiment. Our findings show that, within the framework of the person-artefact-task model, cognitive and experiential factors, together, do indeed influence task outcomes in web navigation. In particular, artefact complexity and task complexity have an effect on task performance, flow and task outcome (*Hypotheses 1-2*). Flow is a partial mediator of the effect of site- and task complexity on task performance (*Hypothesis 5*). Task performance is a complete mediator of the effect of flow on task outcome (*Hypotheses 4 and 6*).

Hypotheses 1 and 2. The effect of artefact complexity (*Hypothesis 1*) on task performance and task outcome is consistent with Gwidzka and Spence's (2006) idea that task difficulty increases with the number of navigation choices and with Blackmon et al.'s (2002) finding of lower success rate with increased number of links per page. Furthermore, the dual-criterion model of menu selection (Pierce et al., 1992) is consistent with our results. The effect of task complexity (*Hypothesis 2*) on task performance and task outcome is consistent with van Oostendorp et al.'s (2009) findings and with the mechanisms derived from Gwizdka and Spence (2006). Its effect on flow is consistent with the idea – based on Guo & Poole (2009) – that as task complexity increases the balance between challenge and skill is adversely affected and flow experience decreases.

The effect of task complexity was stronger than artefact complexity in terms of effect size. In the domain of web navigation, the effect of task complexity is based on the idea that, irrespective of site complexity, the chance of errors in link selection and relevance judgement multiplies with the number of pages visited. In this domain, the effect of artefact complexity, however, is based on the idea that, irrespective of task complexity, the chance of errors in link selection increases with the number of links per page. Assuming that the error rate of selection and relevance judgement is independent across pages, the error rate would increase exponentially with task complexity, that is the (minimum) number of pages needed to complete a task. In relation to artefact complexity, the plausibility of distractors is likely to increase with the number of items, consistent with the dual-criterion model of menu selection (Pierce et al., 1992). Therefore, the more links on a page the higher the error rate. The form of the (monotonic) function between error rate and number of links is not certain at this stage, but – given the findings reported here and the range of artefact complexity studied – the function would be less steep than that for task complexity. Furthermore, task complexity moderated the effect of artefact complexity on correctness of completed tasks. This finding would be consistent with a steeper function for task complexity, but only if artefact complexity is sufficiently high.

In sum, several findings support the conclusion that both artefact- and task complexity have an effect on human-computer interaction with information-oriented Web sites. However, in the domain of web navigation the effect of task complexity seems stronger than that of artefact complexity.

Hypothesis 3. The effect of intrinsic motivation on flow was not significant. It appears that the composite measure of flow was not sufficiently specific to capture the effect of intrinsic motivation and therefore the effect of intrinsic motivation on task performance and task outcome, which would be mediated by flow, was not significant either. In order to explore why intrinsic motivation did not have a significant effect on flow, correlations with individual dimensions of flow were examined (see Section 4.1 and Figure 5). There was a notable correlation of intrinsic motivation with autotelic experience ($r = 0.49$), which remained significant after controlling for task complexity and artefact complexity. This result is consistent with (a) the concept of the autotelic personality: intrinsically motivated individuals should experience a higher level of flow than others because they engage in activities for the sake of the activities rather than in order to achieve some external goal and (b) Asakawa's (2004) findings. According to Nakamura and Csikszentmihalyi (2002), autotelic individuals have higher-order skills which allow them to achieve and sustain (a high level of) flow. These skills include the ability "to continuously adjust the balance between perceived challenges and skills" (Asakawa, 2004, pp. 149-150). However, in the current study, participants were not given the opportunity to apply this skill, as the set of tasks was fixed. This is possibly why intrinsic motivation did not have an effect on flow (as an aggregate), task performance and task outcome beyond its effect on autotelic experience (as a dimension of flow).

In sum, there was no empirical evidence to support the conclusion that intrinsic motivation has an effect on flow as a composite measure, but intrinsic motivation was an independent positive predictor of autotelic experience. Limitations regarding the generality of any effect of intrinsic motivation are discussed in the next section.

Hypotheses 4, 5 and *6*. First, it had to be established that flow is a mediator of the effect of the experimental manipulations of artefact complexity and task complexity on task performance. The results of mediation analysis demonstrate that flow is indeed such a (partial) mediator. Second, it had to be established that task performance is a mediator of the effect of flow on task outcome. The results of mediation analysis confirmed task performance as such a (full) mediator. Thus *Hypotheses 4, 5* and *6* were all confirmed. The results for *Hypothesis 4* (flow has a positive effect on task outcome) are consistent with the results of previous research in other domains (Engeser & Rheinberg, 2008; Heine, 1997; Vollmeyer & Imhof, 2007), demonstrating the positive effect of flow on task outcome, supported by two pathways proposed by Engeser and Rheinberg (2008). The results for *Hypothesis 5* (flow has a positive effect on task performance) and *Hypothesis 6* (task performance has a positive effect on task outcome) confirm the role of task performance as a mediator of the effect of flow on task outcome.

In sum, the findings support the conclusion that flow is a mediator of the effect of the experimental manipulations of artefact complexity and task complexity on task performance. Furthermore, task performance is a mediator of the effect of flow on task outcome.

5.2. Implications of the research model and future work

We explore further implications from the research model in the light of theoretical and empirical work reported in the literature. The three components of the person-artefact- task model are each considered in turn, as well as the role of flow.

Person. Steele and Fullagar (2008) found that perceived need for autonomy is a moderator of the effect of intrinsic motivation on flow. Therefore, the effect of intrinsic motivation on flow dimensions (*Hypothesis 3* in the current study) could not be fully accounted for and was possibly underestimated in the current study, as perceived need for autonomy was not included. Thus, a goal of future research would be to elucidate the generality of the effect of this moderator in Web navigation. However, irrespective of the moderator, our findings do demonstrate the effect of intrinsic motivation on the dimension of autotelic experience. Research has shown that variables other than intrinsic motivation, in particular cognitive individual-difference variables, have an effect on web navigation. Examples of these variables include spatial ability (Juvina & van Oostendorp, 2006; Blustein et al., 2009; Gwizdka, 2009), working-memory capacity (Juvina & van Oostendorp, 2006; Gwizdka, 2009), domain expertise (Juvina & van Oostendorp, 2006) and cognitive style (Gwizdka, 2008). A question for future research is then: to what extent do personality variables (such as intrinsic motivation) and cognitive variables independently influence experiential and cognitive outcomes (flow, task performance and task outcome)?

Artefact. Two major artefact characteristics that have been studied in terms of their consequences for the quality of human-computer interaction are usability and aesthetics. For example, Moshagen et al. (2009) found that – under conditions of poor usability – enhanced (design of) *aesthetics* can improve task outcome. The mediator(s) of the effect of aesthetics, both cognitive (e.g. quality of information-finding performance) and experiential (e.g. flow), remain to be established. Generally, *usability* has a positive effect on task outcome. Moreover, the findings of Pace (2004) and Pilke (2004) demonstrate a remarkable similarity in the antecedents of flow, task performance and task outcome: the characteristics of a usable user-interface appear to be the same as those of a flow-inducing interface! These include transparency of the interface, immediate feedback and (design of) aesthetics. In a theoretical analysis, one type of aesthetics, classical aesthetics, has since been identified as being correlated with usability (van Schaik & Ling, 2009). Therefore, first, as demonstrated in the current study, flow is a mediator of the effect of task complexity and artefact complexity on task performance, so flow would be a mediator of the effect of the usability of system design on task performance. Second, as demonstrated in the current study, task performance is a mediator of the effect of flow on task outcome while controlling for task- and artefact complexity, so task performance could be a mediator while controlling for usability. The correctness of both of these conjectures would have to be established in future research.

An important complex of artefact characteristics to be addressed in future research involved is the information architecture of web sites (see also Katz & Byrne, 2003; Miller & Remington, 2004; Pak & Price, 2008; Resnick & Baker, 2009; Resnick & Sanchez, 2004). The three major components are conceptual (organisational) structure, navigation structure and labelling system (Rosenfeld & Morville, 2006), but other artefacts characteristics such as layout and use of colour are not part of their information architecture. The first is the organisation of larger (clusters of) content elements (e.g. text or graphics) into the web pages comprising a web site and the organisation of individual content elements under headings within a page. The second is the linking of content elements across a web site using hyperlinks. The third is the language used to describe (‘label’) hyperlinks and headings. Empirical psychological techniques like card sorting and rating have been used to aid in the design of the information architecture of small web sites. However, the use of these user-based techniques and other expert-based techniques does not scale up because of the immense cognitive complexity and the massive amount of time required when applied to large web sites. Therefore, for large sites another approach is required and computational cognitive modelling techniques, with a cognitive-psychological basis, should be explored as they would be expected to scale up and would not be constrained by limited human capabilities in terms of information processing and limited availability of time. There is relevant research into design support for the conceptual structure and, to some extent, the navigation structure (Katsanos, Tselios & Avouris, 2008). There is also research on the cognitive computational modelling of web navigation and – based on this modelling – 'repairing' usability problems (e.g. Blackmon et al., 2005), and providing online support (van Oostendorp & Juvina, 2007). However, there is a lack of work explicitly modelling (a) an information architecture in its entirety and (b) in relation to web navigation. Research addressing these issues should focus on explicitly modelling information architecture, database modelling to reflect a cognitive model of web navigation, batch-mode automation of the modelling of web navigation (which will be essential to support the comprehensive simulation of users' navigation of large web sites) and automated support for the improvement of information architectures and – thereby – web sites. This focus should make it possible to support the design of large web sites and the cognitive-experiential modelling of web navigation. The modelling of web navigation should also make it possible to simulate the effect of different information architectures modelled as a database. The best information architecture according to simulation results would then be selected to create an improved web site. Information architecture modelled this way would then be used to (semi-)automatically create a database-driven web site. More generally, a strength of this approach would be the global/sitewide analysis of information architecture and web navigation in order to avoid 'local optimisation' of the architecture to the detriment of other parts of the architecture. The modelling would be conducted within the person-artefact-task framework in order to be able to provide a comprehensive account of web navigation.

Task. Two important task characteristics that are relevant to the study of human-computer interaction are mode of use and skill requirement. Hassenzahl (e.g. Hassenzahl & Ullrich, 2007) has introduced the concept of *mode of use* to describe the mental state of a user in relation to a product or system. According to Hassenzahl (2003, p. 39), ‘‘usage *always* [emphasis in original] consists of behavioural goals and actions to fulfil these goals’’. When human-system interaction occurs in goal mode, users focus on the accomplishment of goals and the product is just ‘a means to an end’ (Hassenzahl, 2003, p. 39). In action mode, users focus on actions and the product can be ‘an end in itself’ (Hassenzahl, 2003, p. 40) and can thus be seen as driven by intrinsic motivation. Van Schaik and Ling (2009) found that when a mode of use is provided, test-users’ ratings of aesthetics become more consistent over time. Therefore, mode of use can be considered to provide users with a context for their interaction with an artefact. The particular mode of use is triggered by the situation in which the product is used. Therefore, for tasks in goal mode (as in the current study) flow would be a determinant of task performance and task outcome. In action mode, the same would be true in principle, but more important is that in this mode flow would be a goal in its own right and, given its nature (driven by intrinsic motivation), it would be characterised by a high level of flow.

Regarding skill requirement for task performance, a basic distinction is made between task performance that is based on recurrent skills and that based on non-recurrent skills (van Merriënboer, 1997). Recurrent skills are rule-based, highly domain-specific, algorithmic procedures that guarantee that task goals will be reached. Non-recurrent skills consist of higher-level, strategic knowledge in the form of cognitive schemas used to solve unfamiliar aspects of problems; these schemas can be effective, but do not guarantee that task goals will be achieved. In reality, many tasks are based on a combination of both recurrent and non-recurrent skills. Tasks will then be more difficult the more skills they require to execute and co-ordinate (but obviously this also depends on a person's existing skills and prerequisite knowledge). In terms of task difficulty, the effect of skill requirement is likely to be similar to that of usability, with lower task difficulty being beneficial for flow, task performance and task outcome. However, a difference between level of skill requirement and usability is that a high level of usability is normally considered to be desirable whereas the appropriate skill requirement for task performance depends on task goal and a user's existing skills and prerequisite knowledge and, therefore, cannot be controlled by system design. Rather, system design should explicitly support tasks with different skill requirements with appropriate functionality and high usability in order to promote flow, task performance and task outcome. This is an aim of electronic performance support systems (Barker & van Schaik, 2010). For example, a web application that explicitly supports the comparison of products on relevant product dimensions would support decision-making in the purchase of goods.

Flow. In the framework of need fulfilment in human-computer interaction, flow can be 'understood as a variant of a competence experience' (Hassenzahl et al., 2010). This view allows a more integrated analysis of flow as a multiple antecedent: not only of cognitive task performance (as studied here), but also of positive affect and 'product' (artefact) perceptions within Hassenzahl's interaction experience model. Two observations are of critical importance. First, from this perspective, the current study meets the requirement – identified by Sheldon et al. (2001) – to validate need fulfilment (of which flow is an example in Hassenzahl's conceptualisation) with objective measures, in this case task performance and task outcome. Second, as argued theoretically and confirmed empirically by Hassenzahl et al., the positive effect of need fulfilment depends on/is moderated by 'attribution'. This means that the extent to which a user attributes the flow that is experienced to the artefact that is being used influences the effect of need fulfilment: the stronger the attribution, the stronger the effect. These conjectures and this more integrated view of flow as a multiple antecedent of cognitive task performance and product perceptions should be investigated in future research.

The finding that specific dimensions of flow combine to have a positive effect on task performance and task outcome lead to proposals for enhancing these two types of outcome by way of enhancing flow. In particular, based on our results, characteristics of artefacts and tasks that enhance the mergence of action and awareness, clarity of goals, balance of challenge and skills, control and feedback would indirectly enhance these outcomes mediated by flow. Future research should demonstrate the validity of this conjecture experimentally.

It is also important to note that, first, although flow experience was theoretically and empirically found to be an independent positive predictor of task outcome, the current study did not control for other artefact characteristics variables (such as aesthetics and [other] pleasure-producing qualities). Second, although high task complexity and goal frustration are directly related to emotional responses with possible effects on task performance, test users' affective response was not measured. These two issues would be the subject of future research in order to broaden our understanding of the extent of the role of flow in human-computer interaction.

6. Conclusion

Based on theoretical and empirical grounds, we have argued the need for an integrated approach to studying cognitive and experiential factors in human-computer interaction in order to model web navigation, based on the person-artefact-task framework. Our study provides empirical evidence for this approach. In addition to a more complete theoretical understanding of processes of human-computer interaction, another benefit of this approach is the promise of a greater amount of variance explained in task outcome. We are therefore looking forward to seeing this approach fruitfully applied to this field of research and its beneficial consequences realised.

Acknowledgement

The authors are grateful to three anonymous reviewers for their helpful comments on the first draft of the manuscript and to the editor Andrew Monk for superb guidance.

Notes

Authors' Present Addresses. Paul van Schaik, School of Social Sciences and Law, Teesside University, Borough Road, Middlesbrough, Cleveland, TS1 3BA, United Kingdom. Jonathan Ling, Faculty of Applied Sciences, The Sciences Complex, Wharncliffe Street, University of Sunderland, Sunderland SR1 3SD, UK

References

Ahuja, J., & Webster, J. (2001). Perceived disorientation: An examination of a new measure to assess web design effectiveness. *Interacting with Computers*, *14*, 5-29.

Asakawa, K. (2004). Flow experience and autotelic personality in Japanese college students: How do they experience challenges in daily life? *Journal of Happiness Studies*, *5*, 123-154.

Asakawa, K. (2009). Flow experience, culture, and well-being: how do autotelic Japanese college students feel, behave, and think in their daily lives? *Journal of Happiness Studies*, *9*, 1-19.

Ayers, B., & Forshaw, M. (2010). An interpretative phenomenological analysis of the psychological ramifications of hand-arm vibration syndrome. *Journal of Health Psychology*, *15*, 533-542.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs: Prentice-Hall.

Barker, P., & Schaik, P. van (Eds.) (2010). *Electronic performance support: Using technology to enhance human performance*. Aldershot, Hants: Gower.

Blackmon, M. H., Polson, P. G., Kitajima, M., & Lewis, C. (2002) Cognitive Walkthrough for the Web. *Proceedings of CHI 2002* (pp. 463-470). New York: ACM Press.

Blackmon, M.H., Kitajima, M., & Polson, P.G. (2005). Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs*. Proceedings of CHI 2005 (pp. 31-40). New York: ACM Press.*

Blustein, J., Ahmed, I., Parvaiz, H., Fu, C.-L., Wang, C., Chapman, A., & Hu, Y. (2009). Impact of spatial visualization aptitude on WWW navigation. *The Ergonomics Open Journal*, *2*, 80-87.

Broadbent, D.E. (1958). *Perception and communication*. London: Pergamon Press.

Card, S., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conway, M. (2002). Sensory-perceptual episodic memory and its context: Autobiographical memory. In A. Baddeley, J. Aggleton, & M. Conway (Eds.), *Episodic memory: New directions in research* (pp. 53-70). New York, NY: Oxford University Press.

Creswell, J. (2007). *Qualitative enquiry and research design* (2nd ed.). London: Sage.

Csikszentmihalyi, M. (1990). *Flow: the psychology of optimal experience*. New York: Harper & Row.

Dalal, N.P., Quible, Z., & Wyatt, K. (2000). Cognitive design of home pages: An experimental study of comprehension on the World Wide Web. *Information Processing and Management*, *36*, 607-621.

Damasio, A. (2000). *The feeling of what happens: Body, emotion and the making of consciousness*. London: Vintage.

David, P., Song, M., Hayes, A., & Fredin, E.S. (2007). A cyclic model of information seeking in hyperlinked environments: The role of goals, self-efficacy, and intrinsic motivation. *International Journal of Human-Computer Studies*, *65*, 170-182.

Davis, S., & Wiedenbeck, S. (2001). The mediating effects of intrinsic motivation, ease of use and usefulness perceptions on performance in first-time and subsequent computer users. *Interacting with Computers*, *13*, 549–580.

Engeser, F., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, *32*, 158-172.

Finneran, C., & Zhang, P. (2003). A Person-Artefact-Task (PAT) model of flow antecedents in computer-mediated environments. *International Journal of Human-Computer Studies*, *59*, 475-496.

Gardiner, M.M., & Christie, B. (Eds.) (1987). *Applying cognitive psychology to user-interface design*. Chichester: Wiley.

Gefen, D., Karahanna, E., & Straub, D. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly: Management Information Systems*, *27*, 51-90.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (SIMS). *Motivation and Emotion*, *24*, 175-213.

Guo, Y.M., & Poole, M.S. (2009). Antecedents of flow in online shopping: A test of alternative models. *Information Systems Journal*, *19*, 369-390.

Gwizdka, J. (2008). Revisiting search task difficulty: Behavioral and individual difference measures. *Proceedings of the 71st Annual Meeting of the American Society for Information Science and Technology (ASIST 2008)*. Silver Spring, Maryland: American Society for Information Science and Technology.

Gwizdka, J. (2009). Assessing cognitive load on Web search tasks. *The Ergonomics Open Journal*, *2*, 114-123.

Gwizdka, J., Spence, I., 2006. What can searching behavior tell us about the difficulty of information tasks? A study of web navigation. *Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology (ASIST 2006)*. Silver Spring, Maryland: American Society for Information Science and Technology.

Hassenzahl, M., (2003). The thing and I: Understanding the relationship between user and product. In M. Blythe, C. Overbeeke, A. Monk, & P. Wright (Eds.), *Funology: from usability to enjoyment* (pp. 31-42). Dordrecht, Kluwer.

Hassenzahl, M. (2004). The interplay of beauty goodness and usability in interactive products. *Human-Computer Interaction*, *19*, 319-349.

Hassenzahl, M. (2008). Aesthetics in interactive products: correlates and consequences of beauty. In Schifferstein, H. & Hekkert, P. (Eds.), *Product experience* (pp. 287-302). Amsterdam: Elsevier.

Hassenzahl, M., Diefenbach, S., & Göritz, A. (2010). Needs, affect, interactive products – facets of user experience. *Interacting with Computers*, *22*, 353-362.

Hassenzahl, M., & Monk, A. (in press). The inference of perceived usability from beauty. *Human-Computer Interaction*.

Hassenzahl, M., & Tractinsky, N. (2006). User experience – a research agenda. *Behaviour and Information Technology*, *25*, 91-97.

Hassenzahl, M., & Ullrich, D. (2007). To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers*, *19*, 429-437.

Heine, C.A. (1997). *Task enjoyment and mathematical achievement*. Unpublished dissertation. University of Chicago, USA.

Henseler, J., Ringle, C., & Sinkovics, R. (2009). The use of partial least squares modeling in international marketing. In T. Cavusgil, R. Sinkovics & P. Ghauri (Eds.), *New challenges in international marketing (advances in international marketing)*: Vol. 20 (pp. 277-319). London: Emerald.

Hirschfeld, R.R., Thomas, C.H., & McNatt, D.B. (2008). Implications of self-deception for self-reported intrinsic and extrinsic motivational dispositions and actual learning performance: A higher order structural model. *Educational and Psychological Measurement*, *68*, 154-173.

Hulland, J., Ryan, M., & Rayner, R. (2010). Modeling customer satisfaction: A comparative evaluation of covariance structure analysis versus partial least squares. In V.E. Vinzi, W. Chin, J. Henseler & H. Wang (Eds), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (pp. 307-325). Berlin: Springer.

Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of Sport & Exercise Psychology,* 18, 17-35.

Juvina, I., & Oostendorp, H. van (2006). Individual differences and behavioral metrics involved in modeling web navigation. *Universal Access in Information Society*, *4*, 258-269.

Katsanos, C., Tselios, N., & Avouris, N. (2008). Automated semantic elaboration of web site information architecture. *Interacting with Computers*, *20*, 535-544.

Katz, M., & Byrne, M. (2003). Effects of scent and breadth on use of site-specific search on e-commerce web sites. *ACM Transactions on Computer-Human Interaction*, *10*, 198-220.

Keller, J., & Bless, H. (2008). Flow and regulatory compatibility: An experimental approach to the flow model of intrinsic motivation. *Personality and Social Psychology Bulletin*, *34*, 196-209.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Law, E., & Schaik, P. van (2010, in press). Modelling user experience: A research agenda. *Interacting with Computers*, *22*, 313-322.

McCarthy, J. & Wright, P. (2004). *Technology as experience*. Cambridge, MA: MIT Press.

Melzack, R., & Wall, P. (1965). Pain mechanisms: A new theory. *Science*, *150*, 971-979.

Merriënboer, J.J.G. van (1997). *Training complex skills: A four-component instructional design model for technical training*. Englewood Cliffs, NJ: Educational Technology Publications.

Miller, C.S., & Remington, R.W. (2004). Modeling information navigation: Implications for information architecture. *Human-Computer Interaction*, *19*, 225-271.

Mosenthal, P. (1996). Understanding the strategies of document literacy and their conditions of use. *Journal of Educational Psychology*, *88*, 314-332.

Moshagen, M., Musch, J., & Göritz, A. (2009). A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance. *Ergonomics*, *52*, 1311-1320.

Murhpy, L., Smith, K., & Hancock, P. (2008). An hedonomic evaluation of the effect of repeated system-exposure on pleasurable human-system experience. *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*, 518-522.

Nakamura, J., & M. Csikszentmihalyi (2002). The concept of flow. In C.R. Snyder and S.J. Lopez (Eds.), *Handbook of positive psychology* (pp. 89-105). New York: Oxford University Press.

Norman, D. (2004). *Emotional design*. New York: Basic Books.

Novak, T.P., Hoffman, D.L., & Duhachek, A. (2003). The influence of goal-directed and experiential activities on online flow activities. *Journal of Consumer Psychology*, *13*, 3-16.

Oh, S.-Y. (2001). *Flow in golf: motivation, goal orientation, and challenge determinant*. Unpublished dissertation. West Virginia University, USA.

Oostendorp, H. van, & Juvina, I. (2007). Using a cognitive model to generate web navigation support. *International Journal of Human Computer Studies*, *65*, 887-897.

Oostendorp, H. van, Madrid, R.I., & Puerta Melguizo, M.C. (2009). The effect of menu type and task complexity on information retrieval performance. *The Ergonomics Open Journal*, *2*, 64-71.

Paap, K., & Cooke, N. (1997). Design of menus. In W. Helander, T. Landauer & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction* (2nd ed., pp. 533-572). Amsterdam: Elsevier.

Pace, S. (2004). A grounded theory of the flow experiences of Web users. *International Journal of Human-Computer Studies*, *60*, 327-363.

Pak, R., & Price, M.M. (2008). Designing an information search interface for younger and older adults. *Human Factors*, *50*, 614-628.

Pearson, R., & Schaik, P. van (2003). The effect of spatial layout and link colour of web pages on performance in a visual search task and an interactive search task. *International Journal of Human-Computer Studies*, *59*, 327-335.

Pierce, B., Sisson, N., & Parkinson, S. (1992). Menu search and selection processes: a quantitative performance model. *International Journal of Man-Machine Studies*, 37, 679-702.

Pilke, E. (2004). Flow experiences in information technology use. *International Journal of Human-Computer Studies*, *61*, 347-357.

Resnick, M., & Baker, A. (2009). Information architecture in enterprise systems and intranets. *Proceedings of the 2009 Industrial Engineering Research Conference*. Miami, Florida, May 30-June 3, 2009..

Resnick, M., & Sanchez, J. (2004) . Effects of organizational scheme and labeling on task performance in product-centered and user-centered retail web sites. *Human Factors*, *46*, 104-117.

Ringle, C., Wende, S., & Will, A. (2005). SmartPLS 2.0 M3. Last accessed from the WWW, 8/12/2009, at *http://www.smartpls.de*.

Rizq, R., & Target, M. (2008). 'Not a little Mickey Mouse thing': how experienced counselling psychologists describe the significance of personal therapy in clinical practice and training. Some results from an interpretative phenomenological analysis. *Counselling Psychology Quarterly*, *21*, 29-48.

Rosenfeld, L., & Morville, P. (2006). *Information architecture for the world wide web: designing large-scale web sites* (3rd ed.). Sebastopol, CA: O'Reilly.

Rouet, J.-F. (2006). *The skills of document use: from text comprehension to web-based learning*. Mahwah, NJ: Erlbaum.

Schaik, P. van, & Ling, J. (2006). The effects of graphical display and screen ratio on information retrieval in web pages. *Computers in Human Behavior*, *22*, 870-884.

Schaik, P. van, & Ling, J. (2008). Modelling user experience with web sites: Usability, hedonic value, beauty and goodness. *Interacting with Computers*, *20*, 419-432.

Schaik, P. van, & Ling, J. (2009). The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human-Computer Studies*, *67*, 79-89.

Schaik, P. van, & Ling, J. (in press). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. *Interacting with Computers*.

Schwartz, S., & Waterman, A. (2006). Changing interests: A longitudinal study of intrinsic motivation for personally salient activities. *Journal of Research in Personality*, *40*, 1119-1136.

Sheldon, K., Elliot, A., Kim, Y., & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology*, *80*, 325-239.

Sonderegger, A., & Sauer, J. (2010). The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied Ergonomics*, *41*, 403-410.

Stavrou, N.A. (2008). Intrinsic motivation, extrinsic motivation and amotivation: Examining self-determination theory from flow theory perspective. In F. M. Olsson (Ed.), *New developments in the psychology of motivation* (pp. 1-24). Hauppauge, NY: Nova Science Publishers.

Steele, C., & Fullagar, M. (2008). Motivation and flow: toward an understanding of the dynamics of the relation in architecture students. *The Journal of Psychology*, *142*, 533-553.

Streukens, S. Wetzels, M, Daryanto, A., & Ruyter, K. de (2010). Analyzing factorial data using PLS: Application in an online complaining context. In V.E. Vinzi, W. Chin,

J. Henseler & H. Wang (Eds), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (pp. 567-587). Berlin: Springer.

Sutcliffe, A. (2010). *Designing for user engagement: aesthetic and attractive user interfaces*. San Rafael, CA: Morgan Claypool.

Thaler, R., & Sunstein, C. (2009). *Nudge: Improving decisions about health, wealth and happiness*. London: Penguin.

Usher, M., Olami, Z., & McClelland, J.L. (2002). Hick's law in a stochastic race model with speed-accuracy tradeoff. *Journal of Mathematical Psychology*, *46*, 704-715.

Vansteenkiste, M., Timmermans, T., Lens, W., Soenens, B., & van den Broeck, A. (2008). Does extrinsic goal framing enhance extrinsic goal-oriented individuals' learning and performance? An experimental test of the match perspective versus self-determination theory. *Journal of Educational Psychology*, *100*, 387-397.

Vilares, M., Almeida, M., & Coelho, P. (2010). Comparison of likelihood and PLS estimators for structural equation modeling: A simulation with customer satisfaction data. In V.E. Vinzi, W. Chin, J. Henseler & H. Wang (Eds), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (pp. 289-305). Berlin: Springer.

Vinzi, V.E. Chin, W., Henseler, J., & Wang, H. (2010). *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields*. Berlin: Springer.

Vollmeyer, R., & Imhof, M. (2007). Are there gender differences in computer performance? If so, can motivation explain them? *Zeitschrift für Pädagogische Psychologie*, *21*, 251-261.

Zapata-Phelan, C.P., Colquitt, J.A., Scott, B.A., & Livingston, B. (2009). Procedural justice, interactional justice, and task performance: The mediating role of intrinsic motivation. *Organizational Behavior and Human Decision Processes*, *108*, 93-105.

Zijlstra, R., 1993. *Efficiency in work behaviour: a design approach for modern tools*. Delft, Netherlands: Delft University Press.

Appendix – Questionnaire items

Intrinsic motivation (Guay et al., 2000)

Why are you currently engaged in this activity?

IM1	Because I think that this activity is interesting.
IM2	Because I think that this activity is pleasant.
IM3	Because this activity is fun.
IM4	Because I feel good when doing this activity.

Response format: 7-point Likert scale with endpoints ‘Corresponds not all’ and ‘Corresponds exactly’.

Disorientation (Ahuja & Webster, 2001)

DIS1	I felt lost.
DIS2	I felt like I was going around in circles.
DIS3	It was difficult to find a page that I had previously viewed.
DIS4	Navigating between pages was a problem.
DIS5	I didn't know how to get to my desired location.
DIS6	I felt disoriented.
DIS7	After browsing for a while I had no idea where to go next.

Response format: 7-point Likert scale with endpoints ‘never’ and ‘always’.

Flow (Jackson & Marsh, 1996)

Balance of challenge and skill

F1	I was challenged, but I believed my skills would allow me to meet the challenge.
F10	My abilities matched the high challenge of the situation.
F19	I felt I was competent enough to meet the high demands of the situation.
F28	The challenge and my skills were at an equally high level.

Mergence of action and awareness

F2	I made the correct movements without thinking about trying to do so.
F11	Things just seemed to be happening automatically.
F20	I performed automatically.
F29	I did things spontaneously and automatically without having to think.

Clarity of goals

F3	I knew clearly what I wanted to do.
F12	I had a strong sense of what I wanted to do.
F21	I knew what I wanted to achieve.
F30	My goals were clearly defined.

Feedback

F4	It was really clear to me that I was doing well.
F13	I was aware of how well I was performing.
F22	I had a good idea while I was performing about how well I was doing.
F31	I could tell by the way I was performing how well I was doing.

Concentration

F5	My attention was focused entirely on what I was doing.
F14	It was no effort to keep my mind on what was happening.
F23	I had total concentration.
F32	I was completely focused on the task at hand.

Control

F6	I felt in total control of what I was doing.
F15	I felt like I could control what I was doing.
F24	I had a feeling of total control.
F33	I felt in total control of my body.

Loss of self-consciousness

F7	I was not concerned with what others may have been thinking of me.
F16	I was not worried about my performance during the event.
F25	I was not concerned with how I was presenting myself.
F34	I was not worried about what others may have been thinking of me.

Transformation of time

F8	Time seemed to alter (either slowed down or speeded up).
F17	The way time passed seemed to be different from normal.
F26	It felt like time stopped while I was performing.
F35	At times, it almost seemed like things were happening in slow motion.

Autotelic experience

F9	I really enjoyed the experience.
F18	I loved the feeling of my performance and want to capture it again.
F27	The experience left me feeling great.
F36	I found the experience extremely rewarding.

Response format: 7-point Likert scale with endpoints ‘Strongly agree’ and ‘Strongly disagree’.

Figure captions

Figure 1 Person-Artefact-Task model (adapted from Finneran & Zhang, 2003)

Figure 2 Dimensions of flow experience (Jackson & Marsh, 1996)

Figure 3 Web site versions: (a) complex site and (b) simple site

Figure 4 Coefficients of reliability and convergent validity

[a]Bootstrap, N = 5000

Figure 5 Coefficients of discriminant validity

Note. Off-diagonal values are correlations. Diagonal values are square root of average extracted variance. IM: intrinsic motivation. DIS: disorientation. AA: mergence of action and awareness. AE: autotelic experience. CG: clarity of goals. CS: balance of challenge and skill. CNC: concentration. LS: loss of self-consciousness. CTR: control. TT: transformation of time. F: feedback.

Figure 6 Outcome measures as a function of site complexity and task complexity: (a) performance measures and (b) subjective measures

Note. S: Simple. C: complex. O: overall. ES: effect size (r).

[a]Effect size r for site complexity. [b]Effect size r for task complexity. [c]Effect size r intrinsic motivation.

Figure 7 The effects of artefact- and task complexity, and intrinsic motivation on dependent variables

[a]Bootstrap, $N = 5000$.

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Figure 8 The effects of experimental manipulations, flow and task performance on task outcome

Note. Experimental manipulations: site complexity and task complexity. Task performance: disorientation and work load. Correctness is percentage of tasks completed correctly. Figures in brackets show the results of separate analyses of the effect of experimental manipulations and the effect of flow on task performance (in [b]) and of the effect of task performance on correctness (in [c]).
** $p < 0.01$. *** $p < 0.001$.

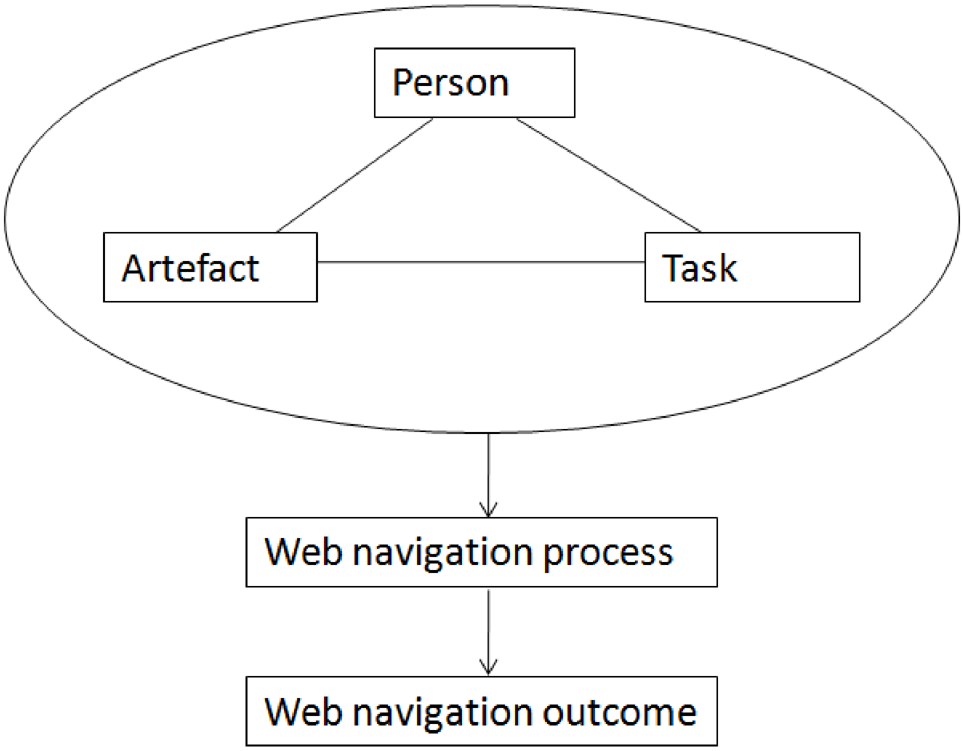


Figure 1 Person-Artefact-Task model (adapted from Finneran & Zhang, 2003)

Dimension	Description
Balance of challenge and skill	"The person perceives a balance between the challenges of a situation and one's skills, with both operating at a personally high level." (p. 18)
Mergence of action and awareness	"The flow activity is so deep that it becomes spontaneous or automatic." (p. 18)
Goal clarity	"Goals in the activity are clearly defined (...), giving the person in flow a strong sense of what he or she is going to do." (p. 19)
Feedback	"Immediate and clear feedback is received, usually from the activity itself, allowing the person to know he or she is succeeding in the set goal." (p. 19)
Concentration	"Total concentration on the task at hand occurs when in flow" (p. 19)
Control	"A sense of exercising control is experienced, without the person actively trying to exert control." (p. 19)
Loss of self-consciousness	"Concern for the self disappears during flow as the person becomes one with the activity." (p. 19)
Transformation of time	"Time alters perceptibly, either slowing down or speeding up" (p. 19)
Autotelic experience	"Intrinsically rewarding experience. An activity is autotelic if it is done for its own sake, with no expectation of some future reward or benefit." (p. 20)

Figure 2 Dimensions of flow experience (Jackson & Marsh, 1996)

a

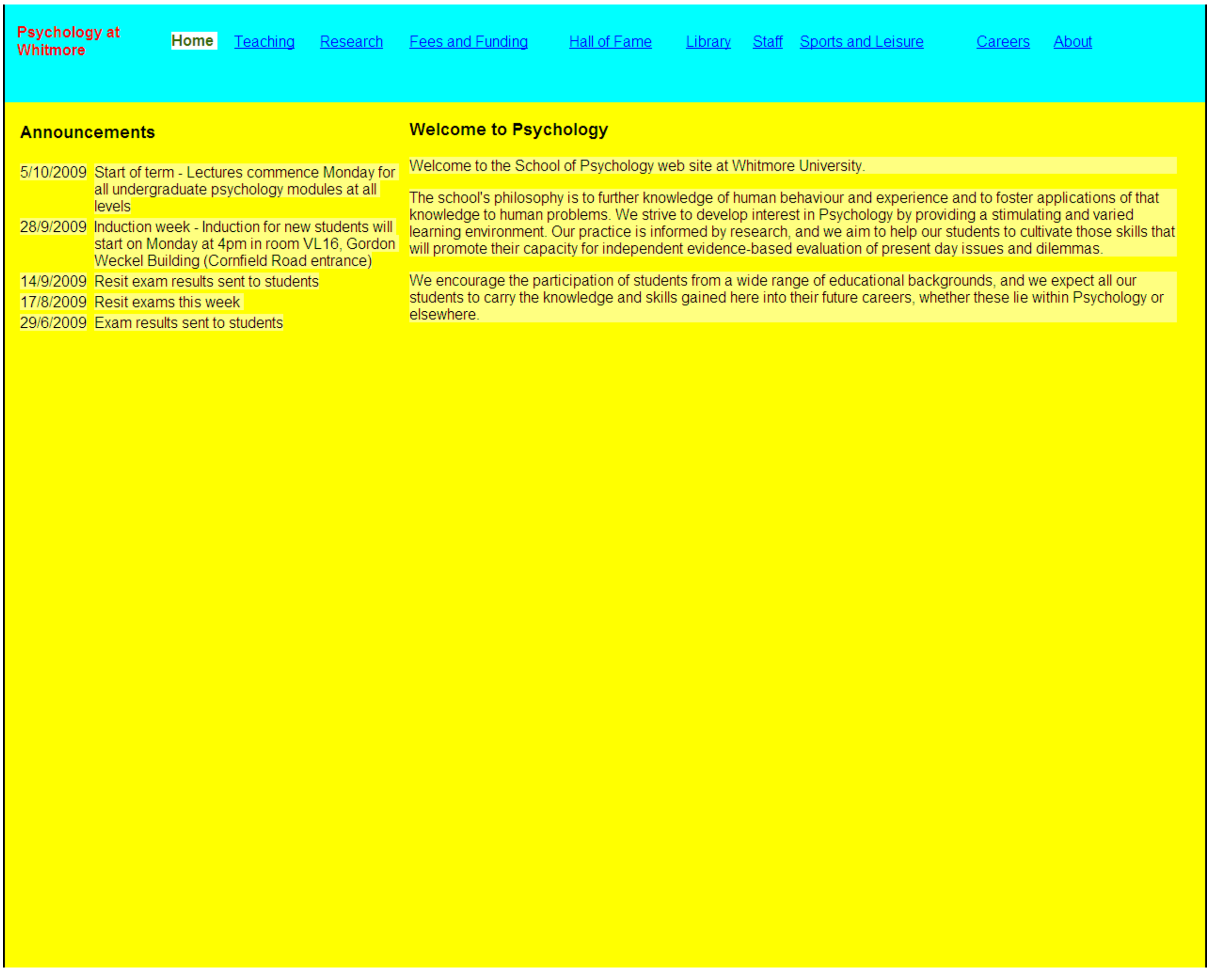


Figure 3 Web site versions: (a) complex site and (b) simple site

b

Figure 3 (continued)

Construct/indicator	Average variance extracted	Composite reliability	Loading	Standard error	t^a
Intrinsic motivation	0.69	0.90			
- IM1			0.76	0.09	8.72
- IM2			0.88	0.05	18.04
- IM3			0.80	0.07	10.93
- IM4			0.88	0.03	25.90
Disorientation	0.74	0.95			
- DIS1			0.86	0.04	19.25
- DIS2			0.90	0.03	34.37
- DIS3			0.78	0.07	11.56
- DIS4			0.79	0.07	11.25
- DIS5			0.89	0.03	29.77
- DIS6			0.90	0.03	31.05
- DIS7			0.88	0.04	20.49
Balance of challenge and skill	0.76	0.90			
- F10			0.88	0.04	21.73
- F19			0.94	0.02	45.50
- F28			0.78	0.08	10.43
Action-awareness merging	0.73	0.92			
- F2			0.83	0.05	17.50
- F11			0.84	0.04	21.71
- F20			0.91	0.03	35.55
- F29			0.85	0.04	20.54
Clarity of goals	0.78	0.93			
- F3			0.90	0.03	32.89
- F12			0.91	0.03	35.08
- F21			0.90	0.04	25.13
- F30			0.82	0.05	15.53
Feedback	0.79	0.94			
- F4			0.86	0.07	12.52
- F13			0.93	0.03	32.80
- F22			0.91	0.03	26.37
- F31			0.86	0.05	18.15
Concentration	0.78	0.93			
- F5			0.89	0.06	14.68
- F14			0.81	0.11	7.49
- F23			0.91	0.05	18.76
- F32			0.92	0.06	15.63

Figure 4 Coefficients of reliability and convergent validity

Control	0.86	0.95			
- F6			0.93	0.02	41.45
- F15			0.95	0.01	68.40
- F24			0.91	0.03	30.31
Loss of self-consciousness	0.72	0.88			
- F7			0.82	0.16	5.23
- F16			0.89	0.13	6.83
- F25			0.84	0.14	6.02
Transformation of time	0.74	0.92			
- F8			0.83	0.09	9.31
- F17			0.87	0.08	10.33
- F26			0.92	0.06	15.63
- F35			0.83	0.09	8.74
Autotelic experience	0.72	0.91			
- F9			0.84	0.04	19.73
- F18			0.81	0.07	11.74
- F27			0.91	0.02	45.27
- F36			0.84	0.04	19.73

Figure 4 (continued)

[a]Bootstrap, N = 5000

	IM	DIS	AA	AE	CG	CS	CNC	LS	CTR	TT	F
Intrinsic motivation	0.83										
Disorientation	0.17	0.86									
Action-awareness merging	-0.07	-0.43	0.86								
Autotelic experience	0.49	0.04	0.17	0.85							
Clarity of goals	-0.17	-0.52	0.70	0.07	0.88						
Balance of challenge and skill	-0.02	-0.46	0.64	0.24	0.56	0.87					
Concentration	0.10	-0.21	0.27	0.33	0.51	0.30	0.88				
Loss of self-consciousness	-0.11	-0.41	0.53	0.00	0.72	0.38	0.49	0.85			
Control	-0.02	-0.61	0.63	0.22	0.72	0.68	0.54	0.59	0.93		
Transformation of time	0.10	0.20	0.08	0.20	-0.17	-0.09	-0.15	-0.10	-0.13	0.86	
Feedback	-0.12	-0.42	0.58	0.24	0.64	0.59	0.44	0.53	0.61	0.02	0.89
Number of tasks completed	-0.23	-0.08	0.41	-0.03	0.18	0.30	-0.13	0.05	0.13	0.10	0.12
Percentage correct	-0.19	-0.29	0.52	0.07	0.36	0.43	0.02	0.19	0.34	0.08	0.26
Percentage completed cor.	-0.05	-0.53	0.45	0.28	0.46	0.49	0.21	0.26	0.57	-0.01	0.44
Time-on-task (correct)	0.21	0.03	-0.14	0.30	-0.02	0.01	0.04	0.00	-0.01	0.05	0.07
Work load	0.04	0.58	-0.57	-0.21	-0.54	-0.54	-0.34	-0.41	-0.68	0.11	-0.52

Figure 5 Coefficients of discriminant validity

	Tasks	Correct	Completed	Time
Percentage correct	0.91			
Percentage completed cor.	0.21	0.52		
Time-on-task (correct)	-0.51	-0.44	0.17	
Work load	-0.24	-0.48	-0.74	0.10

Figure 5 (continued)

Note. Off-diagonal values are correlations. Diagonal values are square root of average extracted variance. IM: intrinsic motivation. DIS: disorientation. AA: mergence of action and awareness. AE: autotelic experience. CG: clarity of goals. CS: balance of challenge and skill. CNC: concentration. LS: loss of self-consciousness. CTR: control. TT: transformation of time. F: feedback.

a

	Number of tasks completed				Percentage correct over all tasks				Percentage completed correct				Time-on-task (correct)			
	Task complexity				Task complexity				Task complexity				Task complexity			
Site complexity	S	C	O	ES	S	C	O	ES	S	C	O	ES	S	C	O	ES
Simple																
- Mean	14.27	7.69	11.03		0.34	0.16	0.25		0.89	0.77	0.83		29.35	49.17	39.09	
- SD	4.33	3.59	5.16		0.1	0.08	0.13		0.12	0.19	0.17		8.74	21.55	19.04	
Complex																
- Mean	11.04	5.7	8.42	0.29[a]	0.25	0.08	0.17	0.4	0.85	0.52	0.69	0.34	49.13	73.61	60.43	0.32
- SD	4.81	4.68	5.42		0.12	0.07	0.13		0.18	0.29	0.29		45.9	45.7	47	
Overall																
- Mean	12.71	6.73	9.77		0.3	0.12	0.21		0.87	0.65	0.76		38.9	60.24	49.09	
- SD	4.81	4.23	5.42		0.12	0.09	0.14		0.15	0.28	0.25		33.71	36.39	36.46	
Effect size	0.57[b]			-0.23[c]	0.68			-0.19	0.49			-0.05	0.32			0.21

Figure 6 Outcome measures as a function of site complexity and task complexity:

(a) performance measures and (b) subjective measures

b

	Work load				Disorientation				Flow			
	Task complexity				Task complexity				Flow complexity			
Site complexity	S	C	O	ES	S	C	O	ES	S	C	O	ES
Simple												
- Mean	16.72	34.35	25.39		1.67	2.68	2.17		5.35	4.59	4.98	
- SD	11.10	22.26	19.49		0.99	1.38	1.29		0.98	1.14	1.12	
Complex												
- Mean	29.98	51.29	40.44	0.32	2.52	3.09	2.80	0.22	5.15	4.31	4.74	0.11
- SD	24.28	31.31	29.71		1.84	1.46	1.67		1.14	1.21	1.24	
Overall												
- Mean	23.12	42.52	32.65		2.08	2.88	2.47		5.26	4.46	4.86	
- SD	19.66	28.08	25.96		1.51	1.42	1.52		1.06	1.17	1.18	
Effect size	0.39			0.03	0.27			0.16		0.34		-0.08

Figure 6 (continued)

Note. S: Simple. C: complex. O: overall. ES: effect size (r).

[a]Effect size r for site complexity. [b]Effect size r for task complexity. [c]Effect size r intrinsic motivation.

a

Dependent variable	Effect	R^2	β	t^a	Significance
Percentage correct (total)		0.52			
	Artefact		-0.30	4.01	***
	Task		-0.63	10.17	***
	Intrinsic motivation		-0.05	0.64	
	Artefact × task		0.02	0.28	
	Artefact × motivation		-0.03	0.45	
	Task × motivation		0.03	0.36	
	Artefact × task × motiv.		-0.05	0.65	
Percentage correct (completed)		0.34			
	Artefact		-0.29	3.71	***
	Task		-0.46	5.84	***
	Intrinsic motivation		0.08	0.75	
	Artefact × task		-0.22	2.72	**
	Artefact × motivation		-0.04	0.40	
	Task × motivation		-0.02	0.17	
	Artefact × task × motiv.		-0.07	0.77	
Number of tasks completed		0.38			
	Artefact		-0.23	2.60	**
	Task		-0.52	6.62	***
	Intrinsic motivation		-0.11	1.26	
	Artefact × task		0.06	0.74	
	Artefact × motivation		-0.03	0.36	
	Task × motivation		0.05	0.58	
	Artefact × task × motiv.		-0.05	0.56	
Time-on-task (correct answers)		0.15			
	Artefact		0.22	2.72	***
	Task		0.21	1.70	
	Intrinsic motivation		0.15	1.41	
	Artefact × task		-0.05	0.46	
	Artefact × motivation		0.07	0.72	
	Task × motivation		0.01	0.12	
	Artefact × task × motiv.		0.05	0.42	

Figure 7 The effects of artefact- and task complexity on (a) performance measures and (b) subjective measures

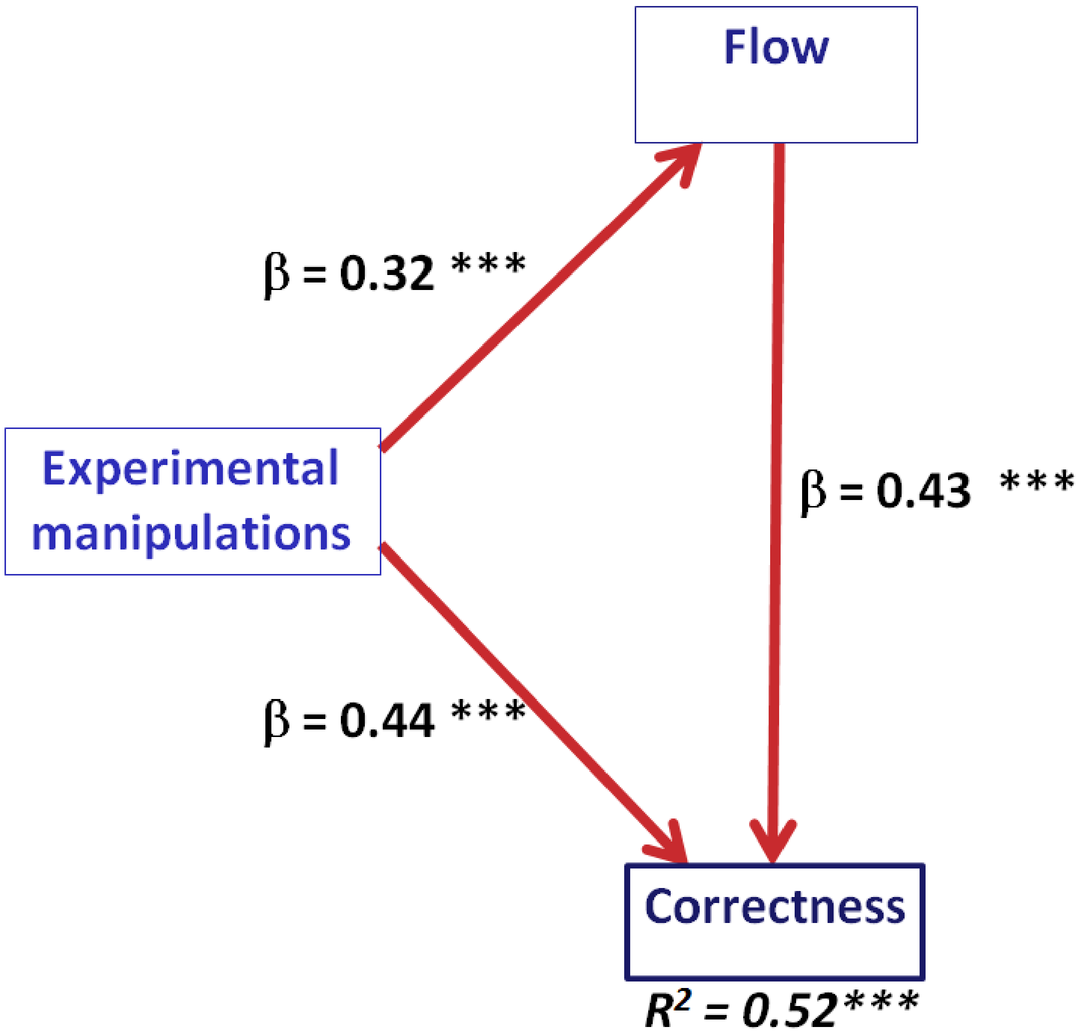
b

Dependent variable	Effect	R^2	β	t^a	Significance
Workload		0.25			
	Artefact		0.30	3.48	***
	Task		0.39	4.33	***
	Intrinsic motivation		-0.03	0.30	
	Artefact × task		0.06	0.64	
	Artefact × motivation		-0.13	1.61	
	Task × motivation		-0.01	0.12	
	Artefact × task × motiv.		-0.04	0.42	
Disorientation		0.15			
	Artefact		0.21	2.06	*
	Task		0.26	2.31	**
	Intrinsic motivation		0.14	1.05	
	Artefact × task		-0.06	0.59	
	Artefact × motivation		-0.06	0.50	
	Task × motivation		-0.06	0.50	
	Artefact × task × motiv.		-0.01	0.07	
Flow		0.16			
	Artefact		-0.12	1.17	
	Task		-0.36	3.79	***
	Intrinsic motivation		-0.04	0.32	
	Artefact × task		-0.04	0.43	
	Artefact × motivation		0.15	1.48	
	Task × motivation		-0.05	0.53	
	Artefact × task × motiv.		0.08	0.74	

Figure 7 (continued)

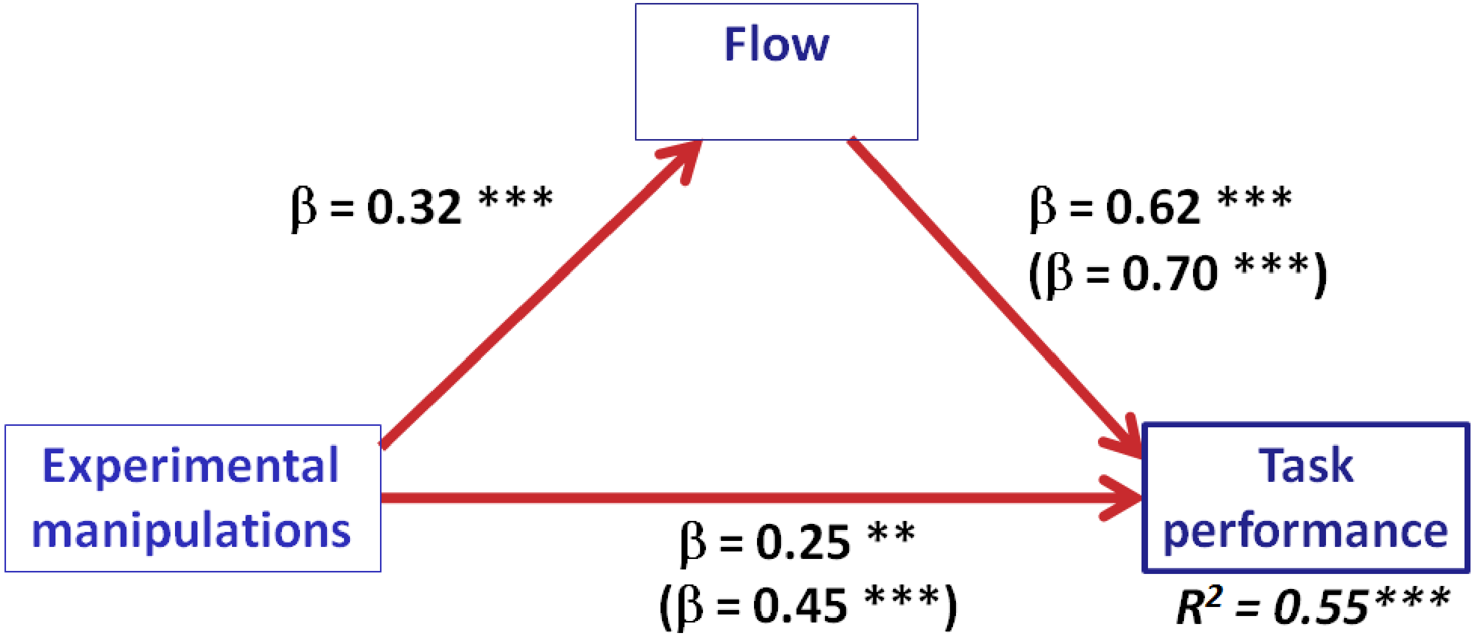
[a]Bootstrap, $N = 5000$.

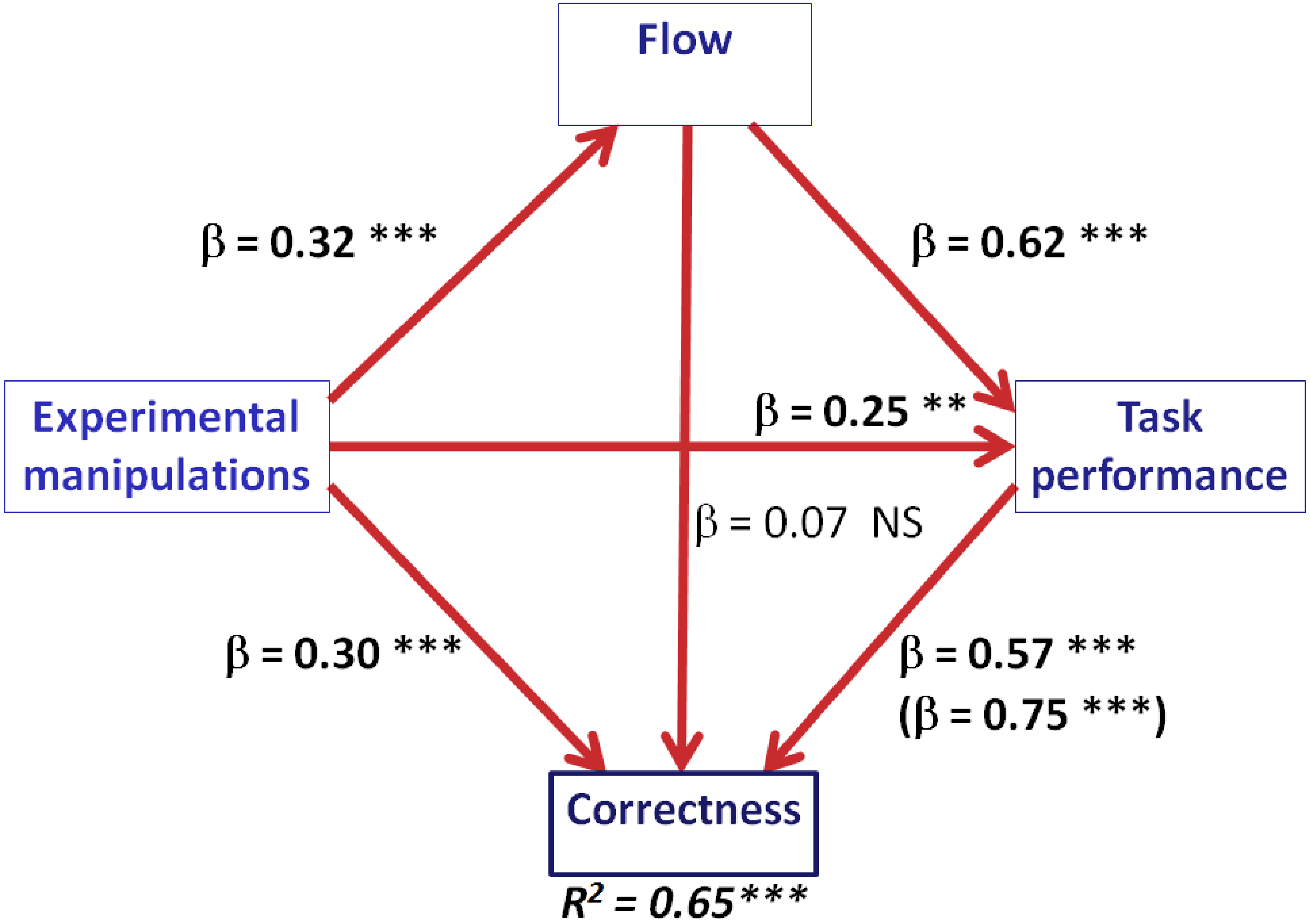
* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.


Flow
Experimental manipulations
Correctness
β = 0.32 ***
β = 0.43 ***
β = 0.44 ***
R2 = 0.52***

a

b


Flow
β = 0.32 ***
β = 0.62 ***
(β = 0.70 ***)
Experimental manipulations
Task performance
β = 0.25 **
(β = 0.45 ***)
R2 = 0.55***


Flow
β = 0.32 ***
β = 0.62 ***
Experimental manipulations
β = 0.25 **
Task performance
β = 0.07 NS
β = 0.30 ***
β = 0.57 ***
(β = 0.75 ***)
Correctness
R² = 0.65***

c

Figure 8 The effects of experimental manipulations, flow and task performance on task outcome

Note. Experimental manipulations: site complexity and task complexity. Task performance: disorientation and work load. Correctness is percentage of tasks completed correctly. Figures in brackets show the results of separate analyses of the effect of experimental manipulations and the effect of flow on task performance (in [b]) and of the effect of task performance on correctness (in [c]).
** $p < 0.01$. *** $p < 0.001$.